Weed Science, 52:936–947. 2004

Exploratory data analysis to identify factors influencing spatial distributions of weed seed banks

L. Wiles
Corresponding author. U.S. Department of Agriculture, Agricultural Research Service, Water Management Research Unit, Fort Collins, CO 80526; lori.wiles@ars.usda.gov

M. Brodahl
U.S. Department of Agriculture, Agricultural Research Service, Water Management Research Unit, Fort Collins, CO 80526

Comparing distributions among fields, species, and management practices will help us understand the spatial dynamics of weed seed banks, but analyzing observational data requires nontraditional statistical methods. We used cluster analysis and classification and regression tree analysis (CART) to investigate factors that influence spatial distributions of seed banks. CART is a method for developing predictive models, but it is also used to explain variation in a response variable from a set of possible explanatory variables. With cluster analysis, we identified patterns of variation with direction of the distance over which seed bank density was correlated (range of spatial dependence) with single-species seed banks in corn. Then we predicted patterns of the seed banks with CART using field and species characteristics and seed bank density as explanatory variables. Patterns differed by magnitude of variation in the range of spatial dependence (strength of anisotropy) and direction of the maximum range. Density and type of irrigation explained the most variation in pattern. Long ranges were associated with large seed banks and stronger anisotropy with furrow than center pivot irrigation. Pattern was also explained by seed size and longevity, characteristics for natural dispersal, species, soil texture, and whether the weed was a grass or broadleaf. Significance of these factors depended on density or type of irrigation, and some patterns were predicted for more than one combination of factors. Dispersal was identified as a primary process of spatial dynamics and pattern varied for seed spread by tillage, wind, or natural dispersal. However, demographic characteristics and density were more important in this research than in previous research. Impact of these factors may have been clearer because interactions were modeled. Lack of data will be the greatest obstacle to using comparative studies and CART to understand the spatial dynamics of weed seed banks.

Nomenclature: Corn, *Zea mays* L.

Key words: Classification and regression tree analysis, seed bank dynamics, spatial dynamics, anisotropy, spatial dependence, geostatistics, spatial correlation, correlograms, seed dispersal, seed longevity, irrigation.

A weed seed bank in a field has spatial pattern if the arrangements of seeds and patches of seeds have a certain amount of predictability so that they can be described quantitatively (Dale 2000). A clear and objective description of spatial pattern can be valuable for generating hypotheses about how biological and environmental factors shape spatial pattern of plant communities (Ford and Renshaw 1984). Identifying factors that create the aggregated spatial pattern of weed seed banks and understanding how patches in seed banks persist or change in size, shape, and density over time could lead to less costly or more accurate methods to sample and map weeds for site-specific weed management (Hausler and Nordmeyer 1999; Rew and Cousens 2001) and the ability to predict the spread of weed populations so that the need for sampling can be minimized (Wilson and Brain 1991). Also, it may be possible to predict the implications of spatially variable management strategies and optimize their effectiveness (Cousens and Woolcock 1997) or identify opportunities to interfere in the spatial dynamics of seed banks to minimize a weed problem (Ghersa and Roush 1993).

Herbicide use may be reduced with site-specific weed management if patchy distributions of weeds in a field can be mapped (Heisel et al. 1999; Johnson et al. 1997; Nordmeyer et al. 1997). The optimal sampling strategy to map a weed population depends on features of the spatial distribution (Burrough 1991; Flatman et al. 1988; Gotway et al. 1996; Oliver et al. 1997; Weisz et al. 1995). For example, observations must be close enough together to be spatially correlated to interpolate a map (Weisz et al. 1995). The relationship between spatial distribution and efficient sampling is significant enough that conducting a preliminary survey to identify some features of a spatial distribution has been recommended before devising a sampling strategy (Burgess et al. 1981). However, surveys to develop prior information about spatial distributions before devising a sampling strategy are unlikely to be cost-effective in most farming systems (Oliver 1999).

If features of the spatial distribution of a seed bank or weed population in a field could be predicted before sampling, there may be information to select a sampling plan that would likely be more efficient than a plan universally recommended for sampling seed banks (Burrough 1991; Flatman et al. 1988; Gotway et al. 1996; Oliver et al. 1997; Weisz et al. 1995). Fewer observations may be needed to make a map, or map accuracy may be improved, if both sample data and the predicted distribution are used to make the map rather than just sample data (Audsley and Beaulah 1996; Hausler and Nordmeyer 1999; Walter et al. 1997). Also, sampling could be targeted to areas where there is the most uncertainty about the management decision (Cardina et al. 1996; Colbach et al. 2000; Dieleman et al. 1999).

Predicting features of the spatial distributions of seed banks requires understanding of the spatial dynamics of seed banks. Empirical research on the spatial dynamics of seed banks has been limited, and the dynamics are poorly understood (Cousens and Mortimer 1995). The dynamics are influenced by management and demographic characteristics of weed species. Dispersal, particularly movement of seeds with a combine harvester, has been identified as a key mechanism of spatial dynamics (Cousens and Woolcock 1997; Paice et al. 1998). Combining may move seeds up to 20 m, and the rate of spread of some weeds may be up to 10 m per year (Howard et al. 1991). In contrast, most seeds are moved 3 m or less by cultivation and at most 1 m by natural dispersal (Grundy et al. 1999; Howard et al. 1991; Mead et al. 1998; Rew and Cussans 1995, 1997). Both long dispersal distances and soil disturbance may lead to less patchy distributions (Paice et al. 1998). Spatial correlation of seed banks has been observed to be stronger in fields managed without tillage compared with fields that were moldboard plowed (Cardina et al. 1996). No clear relationships between demographic characteristics and spatial dynamics have been identified (Cardina et al. 1997; Cousens and Woolcock 1997) except for a few reports of seed size and morphology influencing dispersal by machinery (Grundy et al. 1999; Howard et al. 1991). Rate of population growth is expected to influence spatial dynamics of seed banks (Cousens and Mortimer 1995).

More empirical research and different types of experiments are needed to understand the spatial dynamics of weed seed banks (Cousens and Mortimer 1995). New invasions of weeds in a field have been studied more than established seed banks that occupy most of a field, and cases studies are more common than analyses of the importance of life history and environmental factors in shaping spatial distributions (Cousens and Woolcock 1997). Unfortunately, attempting to identify factors that influence the spatial distribution of a weed population from a single year of data can lead to erroneous conclusions (Cousens and Mortimer 1995). However, observing the spatial distributions of weed seed banks in several fields in a single year could provide useful information if species, management, and soil characteristics vary among fields. Seed banks with similar distributions that may have been shaped by the same set of processes and common management, species, and fields characteristics among those seed banks may indicate the nature and relative importance of the processes.

Investigating the spatial dynamics of seed banks may be more efficient with comparative studies than long-term, controlled studies. Multiple fields can be sampled, and relationships between features of spatial distributions and management and field and species characteristics can be explored. However, traditional statistical analysis of testing predefined hypotheses is rarely appropriate for comparative studies with the incomplete structure of the data (i.e., not every species in every field), lack of independence of distributions of seed banks of different species within a field, and limited prior knowledge of important factors and interactions that determine spatial pattern. However, techniques of exploratory data analysis can identify patterns in observational data that may indicate hypotheses for future research (Hallahan and Rosenthal 2000).

Classification and regression tree analysis (CART) is both an alternative and complement to traditional statistical analysis of ecological data (De'ath and Fabricus 2000). CART is an exploratory data analysis technique for modeling or describing pattern or structure in data sets (Clark and Pregibon 1992). This technique is typically described in terms of using a data set to develop a predictive model for a response variable based on several explanatory variables. However, the predictive model is a description of a structure in the data set and can help us understand underlying causes of the variation in the response variable. CART is well suited for analysis of complex ecological data because the response variable and explanatory variables of the model can be numeric or categorical, data can be unbalanced and contain missing values, and monotonic transformations of explanatory variables do not change the analysis (De'ath and Fabricus 2000). Nonlinear responses and complex interactions can be described, and even complicated structure can be easily interpreted from the tree graph created (De'ath and Fabricus 2000). CART analysis has uncovered structure in ecological data sets that was not detected with analyses using linear models (De'ath and Fabricus 2000; Dobbertin and Biging 1997).

We did a comparative study of spatial pattern of weed seed banks in eight irrigated cornfields and analyzed the data with CART. We had two objectives: (1) to identify correlations between features of the spatial distributions of seed banks and demographic characteristics of species, attributes of fields, and past management; and (2) to investigate the use of CART and analysis of spatial pattern for generating hypotheses about the spatial dynamics of seed banks.

Materials and Methods

We investigated factors influencing the spatial distribution of seed banks of individual weed species with a five-step procedure: (1) collect soil cores in multiple fields and count seeds by species; (2) quantitatively describe spatial pattern of the seed bank of each major species in all fields (geostatistical analysis); (3) identify groups of seed banks with similar spatial patterns (cluster analysis); (4) use CART to develop a model to predict cluster group membership (e.g., seed banks with similar spatial patterns) from weed and field attributes; and (5) examine the structure of the CART model to identify relationships between field and weed attributes and spatial pattern that may help us form hypotheses about the spatial dynamics of seed banks.

Our actual procedure was more complex than the five steps outlined because we analyzed two different features of spatial pattern of the seed banks (cluster analysis was done twice) and examined two different sets of explanatory variables for each feature for a total of four CART models (two different response variables by two sets of explanatory variables).

Seed Bank Sampling

Seed banks were sampled in eight commercial, irrigated cornfields in eastern Colorado. The fields and sampling procedure are described in detail in Wiles and Schweizer (2002). Briefly, seed banks were sampled in an 8.1-ha block of each field in 1993 or 1994, with 1,225 to 1,260 soil cores collected on a square grid. The grid was aligned with the crop row, and sampling locations were 8.4 m apart. Soil cores were 5 cm in diameter and 10 cm deep, seeds were separated from soil using water, and seeds that felt firm after slight squeezing with forceps were counted and identified using a parallel optical stereozoom microscope set at ×7.5 magnification (Wiles et al. 1996). Additional soil cores were collected in a W pattern across each block and composited for texture and organic matter analysis.

Modeling Spatial Correlation

Spatial pattern can be described by attributes such as scale, intensity, and dispersion (Dale 2000). For this analysis, spatial correlation was described. Specifically, we investigated the influence of management, field attributes, and demographic characteristics of weed species on the distance over which seed bank density was correlated and how that distance varied with direction. These characteristics vary among seed banks (Wiles and Schweizer 2002).

Spatial correlation of weed seed bank density was modeled by fitting correlograms to seed count data as described in Wiles and Schweizer (2002). A correlogram describes correlation of seed bank density as a function of distance and direction separating two locations (Isaaks and Srivastava 1989). The value of a correlogram ($\rho(\mathbf{h})$) for distance $\mathbf{h}$ is the correlation coefficient for all pairs of seed counts separated by that distance. The variable $\mathbf{h}$ may be a distance without regard for direction or a vector defined by distance and direction. The range of spatial dependence of a correlogram is the average distance within which observations remain correlated spatially and is estimated by the distance at which the correlogram reaches a plateau (Rossi et al. 1992). Variation in the range of spatial dependence with direction is geometric anisotropy (Deutsch and Journel 1998). Spatial correlation is isotropic if the range of spatial dependence is the same for all directions.

A correlogram was fit for a seed bank (a single species in a field) if seeds of that species were found in 1% or more of the cores from the field. Seed counts were transformed ($\ln(x) + 1$) before analysis to reduce the impact of skewed frequency distributions of seed counts on the spatial analysis. Directional sample correlograms were calculated for the direction of the crop row (0°) and 45, 90, and 135° clockwise from the row. Distances between calculated correlogram values (lag spacings) were matched to the sampling grid with 8.4 m for 0 and 90° clockwise from the crop row and 11.8 m for 45 and 135°. Tolerances for direction and distance determined how many and what pairs of seed counts were included in the calculation of a sample correlogram value (Isaaks and Srivastava 1989). Lag tolerance was 0.5, angular tolerance was 10°, and bandwidth was 17 m. With these specifications, each correlogram value was calculated from at least 730 pairs of seed counts.

Spatial correlation of seed bank density was modeled with sample correlograms rather than the more widely used variograms because correlograms are more robust than variograms for comparing spatial pattern of samples with disparate levels of spatial variability and for describing spatial pattern when local means and variances change within the sample area (Rossi et al. 1992). However, methods and results are presented in variogram form so that parameters could be interpreted as those of that more familiar variogram. A variogram value is 1 minus the correlogram value for the same distance. All sample correlograms were modeled with two nested spherical structures:

$$1 - \rho(\mathbf{h}) = c_0 + c_1 \cdot \mathrm{Sph}_{a_1}(\mathbf{h}) + c_2 \cdot \mathrm{Sph}_{a_2}(\mathbf{h}) \qquad [1]$$

where $\mathbf{h}$ is the separation distance, $\mathrm{Sph}_{a_1}(\mathbf{h})$ and $\mathrm{Sph}_{a_2}(\mathbf{h})$ are spherical models, and c_1 and c_2 are the proportions of the variation in seed bank density explained by each spherical model (Isaaks and Srivastava 1989). The nugget (c_0) is the proportion of the variability that is not modeled as spatial pattern. Spherical models are linear at small lag distances but then flatten out and eventually plateau. The parameter a_i is the distance at which spherical model i reaches a plateau, and the larger of a_1 and a_2 is the range of spatial dependence of the correlogram.

Both sample and theoretical correlograms were calculated using SAGE2001 software[1] (SAGE2001 1999). This software uses regression to fit a correlogram model and simultaneously fits a model to all directional sample correlograms. Consequently, the model of a sample correlogram in one direction is influenced by the nature of sample correlograms in all other directions (SAGE2001 1999). Variation with direction of a_i of a spherical structure is modeled with an ellipse. For any point on an ellipse, a_i is the distance from the center of the ellipse to that point, and direction of a_i is the angle created by the y axis and the vector from the origin to that point. Each ellipse is specified with the distances and directions of the minimum and maximum values of a_i. In some cases, one ellipse may represent the range of spatial dependence for some directions and the second for the remaining directions. Therefore, we calculated the range of spatial dependence for every degree as the maximum of a_i and a_i for that direction to determine the maximum and minimum ranges of the seed bank and the strength of anisotropy. Strength of anisotropy is described as the ratio of maximum range of spatial dependence of a distribution over minimum range (Isaaks and Srivastava 1989).

Cluster Analysis

Each parameter of a correlogram describes a different characteristic of spatial correlation, and all parameters are needed to provide a complete picture of spatial correlation. However, CART can include many explanatory variables, but only one response variable. Therefore, we used cluster analysis to create a single variable, group membership, that incorporated information from all the parameters of a correlogram. Seed banks were grouped with cluster analysis based on a set of ranges of spatial dependence for each seed bank calculated from the values of the parameters of the model of spatial correlation for the seed bank.

Cluster analysis is a set of algorithms for grouping a collection of objects into subsets or "clusters" such that those within a cluster are more closely related to one another than objects assigned to different clusters (Hastie et al. 2001). Objects are judged as similar or dissimilar by a set of attributes. In our analysis, attributes of a seed bank were ranges of spatial dependence calculated for every 15° from 0 to 165° clockwise from the crop row and the algorithm was Ward's minimum variance method (SAS 1988). We used the cluster or group membership as the response variable for the CART analysis. For interpretation of CART results, we de-

Table 1. Field characteristics included as explanatory variables in classification and regression tree analyses of spatial pattern of weed seed banks.

Field	Type of irrigation	Soil texture	% Organic matter	Manure application
C31	Center pivot	Sandy	2.5	Yes
C34	Center pivot	Sandy loam	1.0	No
C42	Center pivot	Sandy clay loam	1.4	No
C44	Center pivot	Sandy clay loam	1.5	No
F32	Furrow	Loam	2.1	No
F33	Furrow	Loam	1.8	No
F41	Furrow	Sandy clay loam	1.8	No
F43	Furrow	Clay loam	2.2	No

scribed the characteristics of the patterns of spatial correlation by calculating the mean and SD of the distance and direction of the maximum range and ratio of anisotropy for each group. We also graphed variation in the range of spatial dependence by direction for all seed banks in each group identified with cluster analysis.

If size of patches for different seed banks that are being compared varies considerably, similarities in the shape of patches may be obscured. For example, the magnitude of the range of spatial dependence may need to be factored out to identify newly introduced seed banks and established seed banks influenced by the same processes. For this reason, we did cluster analysis and CART for both actual and normalized ranges of the seed banks. Ranges were normalized to a maximum of one to investigate anisotropy independent of the distance of the range.

Classification and Regression Tree Analysis

CART models variation of a single categorical or numeric response variable from a set of explanatory variables. Groups of data are repeatedly split on the basis of explanatory variables to create the two most homogeneous groups according to the values of the response variable. The process begins with all the data as the first group, and explanatory variables may be either categorical or numeric (De'ath and Fabricus 2000). A split is defined by a simple rule based on a single explanatory variable (e.g., organic matter is less than 1.5% or soil texture is sandy loam). The two groups resulting from a split are child groups, and terminal groups are child groups that are not split. Splitting is continued until all members of each terminal group have the same value of the response variable or all terminal groups have reached a predefined minimum size. Results of an analysis can be represented with a visual inverted tree-shaped diagram that shows the predicted value of the response variable for each terminal group, all splits used to reach the terminal groups, the rule for each split, and the relative importance of a split in reducing within group variation. When the response variable is categorical, the quality of the model as a whole can be described with a misclassification rate (De'ath and Fabricus 2000). A misclassification is a member of a terminal group that does not have the predicted value of the response variable of that group. The misclassification rate is the proportion of observations that have been misclassified.

CART was done with group membership from cluster analysis of either the actual or normalized ranges as the response variable and demographic characteristics of the weed species and attributes of the field where the sample was collected as explanatory variables. Group membership represents distinct patterns of variation in the range of spatial dependence with direction. Two sets of explanatory variables were used. The first set included type of irrigation, soil texture, percent organic matter, and whether manure was applied to the field (Table 1); dispersal type plus seed size, production, and longevity (Table 2); and seed bank density (Table 3). The second set included all the explanatory variables of the first set plus weed species.

Only general categories were assigned for demographic characteristics because information in the literature was typically limited and inconsistent. Dispersal type was defined as short distance, long distance, or aggregated and long distance to loosely capture the effect of morphological char-

Table 2. Demographic characteristics included as explanatory variables in classification and regression tree analyses of spatial pattern of weed seed banks.

Species[a]	Type of dispersal	Seed longevity	Seed production	Seed size
AMARE	Short distance	Average	High	Small
CCHIN	Long distance	Average	Low	Small
CHEAL	Short distance	Long	High	Small
ECHCG	Short distance	Short	Low	Not small
EPHDE	Short distance	Average	Low	Not small
PANCA	Long distance	Average	Average	Small
POLCO	Long distance	Short	Low	Not small
POROL	Long distance	Long	Average	Small
SETVI	Short distance	Short	Low	Not small
SINAR	Short distance	Long	Low	Small
SOLSA	Aggregated & long distance	Average	Average	Small

[a] Letter code for weed names in WSSA-approved computer code from the *Composite List of Weeds*, Revised 1989. Available only on computer disk from WSSA, 810 East 10th Street, Lawrence, KS 66044-8897.

TABLE 3. Spatial correlation of weed seed banks as described by fitting correlograms and patterns of spatial correlation identified with cluster analysis of actual and normalized ranges of spatial dependence.

Sample[b]	Cores with seeds	Mean seed count[c]	Variability due to spatial pattern[d]	Range of spatial dependence: Maximum	Range of spatial dependence: Direction of maximum range	Range of spatial dependence: Anisotropy ratio[e]	Spatial pattern[a]: Actual ranges	Spatial pattern[a]: Normalized ranges[f]
	%	seeds m^{-2}	%	m	degree from crop row[g]			
C31AMARE	96	4,062 ± 6,961	31	186	1	2.1	I/L[h]	AM/NR[i]
C31CCHPA	1	15 ± 143	35	85	−79	1.2	I/M	I/NR
C31CHEAL	9	138 ± 961	37	154	60	2.7	—	—
C31ECHCG	11	138 ± 1,109	32	148	−40	1.5	I/L	I/NR
C31SETVI	65	991 ± 1,385	41	160	51	1.1	I/L	I/NR
C31SINAR	4	30 ± 168	47	60	0	1.5	A/S	I/NR
C31SOLSA	21	242 ± 744	28	99	−79	1.7	I/M	AM/P
C34AMARE	27	202 ± 518	21	110	10	10.9	—	—
C34CHEAL	4	49 ± 385	33	112	13	10.2	A/S	AS/R
C34POROL	4	30 ± 163	33	115	−65	2.4	I/M	AM/P
C34SOLSA	28	394 ± 1,405	14	80	−62	1.3	I/M	I/NR
C42AMARE	63	1,257 ± 2,509	50	82	−53	1.3	I/M	I/NR
C42ECHCG	1	30 ± 779	48	37	−1	2.2	A/S	AM/NR
C42EPHDE	12	74 ± 360	30	100	−89	1.7	I/M	AM/P
C42PANCA	9	64 ± 266	33	95	3	1.3	I/M	I/NR
C42POLCO	27	325 ± 848	35	144	63	1.1	I/L	I/NR
C42POROL	2	10 ± 99	44	76	−15	4.5	A/S	AS/R
C42SETVI	10	54 ± 182	27	75	−7	1.2	A/M	I/NR
C42SOLSA	35	463 ± 1,242	32	100	3	1.8	A/M	AM/NR
C44AMARE	24	163 ± 370	19	80	0	1.3	A/M	I/NR
C44ECHCG	10	64 ± 232	29	80	13	1.5	A/M	I/NR
C44EPHDE	18	108 ± 247	13	90	−14	1.9	A/M	AM/NR
C44SOLSA	4	30 ± 340	24	60	56	2.4	A/S	AM/NR
F32AMARE	12	74 ± 212	14	152	−1	3.3	A/M	AS/R
F32CHEAL	6	39 ± 168	42	125	0	8.3	A/S	AS/R
F32ECHCG	12	69 ± 217	18	51	−2	3.4	A/S	AS/R
F32SOLSA	13	84 ± 242	22	120	0	2.9	A/M	AS/R
F33EPHDE	3	15 ± 104	48	130	2	3.2	A/M	AS/R
F33SOLSA	14	99 ± 345	27	51	90	1.1	A/S	I/NR
F41AMARE	12	79 ± 251	18	60	−1	3.5	A/S	AS/R
F41ECHCG	3	15 ± 108	26	80	−1	4.7	A/S	AS/R
F41POROL	6	143 ± 1,262	43	100	−90	1.1	I/M	I/NR
F41SETVI	6	35 ± 138	24	60	0	3.2	A/S	AS/R
F41SOLSA	5	35 ± 192	14	32	0	1.3	A/S	I/NR
F43AMARE	7	99 ± 1,114	33	80	−1	5.7	A/S	AS/R
F43POROL	9	158 ± 1,982	71	49	26	2.5	A/S	AM/NR

[a] Seed banks were grouped by similar spatial pattern using cluster analysis.
[b] Sample name indicates field (i.e., C31) and weed species (WSSA code).
[c] Data are the mean count ± SD.
[d] $c_1 + c_2$ of Equation 1.
[e] Maximum range divided by the minimum range.
[f] Ranges scaled to one for the maximum range of spatial dependence of a distribution.
[g] Negative numbers indicate direction counterclockwise from the crop row.
[h] Designations for patterns of variation in actual ranges indicate variation in the range of spatial dependence with direction (anisotropic [A] or isotropic [I]) and length of ranges (short [S], medium [M], or long [L]).
[i] Designations for pattern of variation in normalized ranges indicate variation in the range of spatial dependence and with direction (strong anisotropy [AS], medium anisotropy [AM], or isotropic [I]) and direction of the maximum continuity (row direction [R], nearly row direction [NR], and perpendicular to row direction [P]).

acteristics of a species on dispersal (Table 2). Most species were considered to have mostly upright growth and no significant adaptation for spread of seeds (short distance). Nightshade (primarily *Solanum sarrachoides* Sendtner SOLSA) was separated from other species for the dispersal of seeds in berries resulting in longer distance but more aggregated dispersal of seeds (aggregated and long distance). Purslane (*Portulaca oleracea* L. POROL) and wild buckwheat (*Polygonum convolvus* L. POLCO) were classified as long-distance dispersal for a more horizontal, spreading growth habit; longspine sandbur [*Cenchrus longspinus* (Hack.) Fern. CCHPA] for burs that may be transported over long distances; and witchgrass (*Panicum capillare* L. PANCA) for spread of seeds by mature plants that may break away and tumble. For the second analysis of each type of pattern, species was added as an explanatory variable to evaluate if the demographic variables we selected captured the most important species characteristics influencing the spatial pattern of seed banks. Stopping criteria for all the CARTs were a minimum size of six seed banks for a group to be a candidate for splitting and a minimum child group size of three seed banks (Venables and Ripley 1999).

Results and Discussion

Spatial Correlation

Correlograms were fit for a total of 36 seed banks, with two to eight seed banks of individual species in a field. In all, seven broadleaf and four grass species were observed (Table 3), but not all species were observed in every field. Species, in order of frequency, were pigweed (primarily *Amaranthus retroflexus* L. AMARE) and nightshade in seven fields each, barnyardgrass (*Echinochloa crus-galli* L. ECHGC) in five fields and purslane in four fields. Foxtails [primarily *Setaria viridis* (L.) Beauv. SETVI], common lambsquarters (*Chenopodium album* L. CHEAL) and toothed spurge (*Euphorbia dentata* Michx. EPHDE) were observed in three fields each. The study included only one seed bank each of longspine sandbur, wild buckwheat, witchgrass and wild mustard [*Brassica kaber* (DC.) L. C. Wheeler SINAR]. Observed seed counts were typical for seed banks (Ambrosio et al. 1997; Benoit et al. 1989; Bigwood and Inouye 1988; Chauvel et al. 1989; Dessaint et al. 1996; Jones 1998). That is, frequency distributions of counts of a seed bank were skewed toward low seed counts, variability between counts was high, and the proportion of zero counts was large. Mean seed count of a seed bank ranged from 10 to 4,062 seeds m^{-2}, and seeds were found in 4 to 99% of the cores of a seed bank (Table 3).

Based on our correlogram models, from 13 to 71% of the variability of seed bank density was attributed to spatial pattern ($c_1 + c_2$ of Equation 1) and the average value was 32% (Table 3). Most distributions were anisotropic. The maximum range of spatial dependence of a seed bank (32 to 186 m) was 1.1 to 10.9 times larger than the minimum range (14 to 141 m), and the maximum range was at least twice as long as the minimum range for half of the distributions (Table 3). The direction of maximum continuity (direction of the longest range) was the direction of the crop row (0 ± 3°) for 16 seed banks, perpendicular to the crop row (90 ± 3°) for three seed banks, and 7 to 80° from the crop row for the remaining 17 seed banks. Distances of the

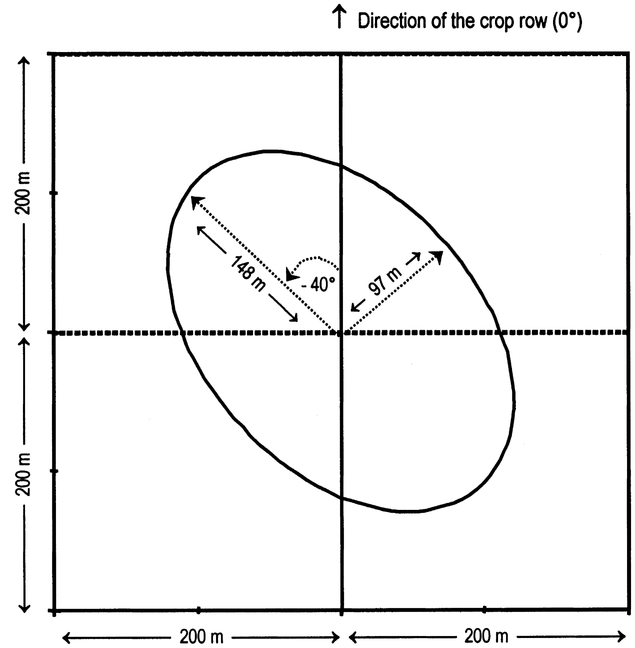

FIGURE 1. Variation in the range of spatial dependence with direction of the barnyardgrass seed bank in Field C3. For any point on the ellipse, the distance from the origin to that point represents the distance of the range of spatial dependence for the direction represented by the angle created by the *y* axis and the vector from the origin to the point.

range of spatial dependence in this study were longer than those observed in studies of common lambsquarters and annual grass seed banks (Cardina et al. 1996) and total seed bank density (Halstead and Gross 1990). However, presence of anisotropy and maximum continuity in the direction of the crop row are consistent with previous research (Benoit et al. 1989).

Spatial Patterns Identified with Cluster Analysis

The common lambsquarters seed bank in Field C31 and the pigweed seed bank in Field C34 were excluded from the cluster analysis because anisotropy of these seed banks had to be modeled as two narrow ellipses oriented in different directions. Spatial correlation of these seed banks was likely shaped by two equally important processes that acted in different directions. For the other seed banks, one ellipse was nearly or completely embedded within the second ellipse. Typically, one ellipse is embedded in another for a seed bank because that ellipse represents spatial correlation from a process acting over short distances (e.g., natural dispersal) and the other ellipse represents spatial correlation from a process acting over longer distances (e.g., spread of seeds by a combine).

For the remaining seed banks, four patterns of the variation in the range of spatial dependence were identified for both actual ranges and another four patterns for normalized ranges. Patterns are shown by plotting the ellipse, or near ellipse, of each seed bank in that group with the *y* axis representing the direction of the crop row (0°) and the *x* axis representing the direction perpendicular to the row (90°) (Figure 1). For any point on an ellipse, the actual or normalized range of spatial dependence is the distance from the origin of the graph to that point. Direction of that range is the angle created by the y axis and the vector from the origin to that point. Positive numbers indicate direction clockwise from the y axis, and negative numbers indicate directions counterclockwise. The ellipse in Figure 1 illustrates the actual ranges of barnyardgrass in Field C31 calculated from the correlogram model for this seed bank. Lines from the origin to a point on the ellipse are shown in the graph for the maximum (148 m) and minimum ranges (97 m). The ratio of anisotropy (maximum range divided by minimum range) of this seed bank was 1.53. Isotropic distributions have ratios of close to 1. With a single ellipse as the model of anisotropy, directions of minimum range and maximum range differ by 90°. Direction of maximum continuity was −40° and direction of the minimum range was 50°.

The four patterns of variation of actual ranges were differentiated by both variation in the range with direction and distances of the range of spatial dependence (Figure 2). The patterns can be generally described as anisotropic with short ranges (A/S), anisotropic with medium ranges (A/M), isotropic with medium ranges (I/M) and isotropic with long ranges (I/L). Pattern A/S (14 seed banks) consisted of seed bank distributions with a maximum range of 67 ± 26 m, whereas pattern I/L (four seed banks) had a maximum range of 160 ± 19 m. Seed banks of patterns A/M (eight seed banks) and I/M (eight seed banks) had intermediate ranges (maximum ranges of 103 ± 28 m and 95 ± 12 m). The ratio of anisotropy of the seed banks ranged from 1.5 ± 0.7 (I/L and I/M) to 3.9 ± 2.7 (A/S). Direction of maximum continuity also varied among the patterns from nearly in the direction of the crop row for all seed banks of pattern A/M (−1 ± 8°) but closer to perpendicular to the row (−64 ± 30°) for seed banks of pattern I/M.

Names of groups identified with cluster analysis of normalized ranges indicate both strength of anisotropy and direction of maximum continuity. The seed banks were separated into three of the four groups by strength of anisotropy: nearly isotropic (I/NR) with a ratio of anisotropy of 1.3 ± 0.2, medium anisotropy (2.1 ± 0.3; AM/NR), and strong anisotropy (4.8 ± 2.4; AS/R) (Figure 3). Three seed banks (AM/P) were grouped separately based on direction of maximum continuity (Figure 3). Direction of maximum continuity was nearly the direction of the crop row (−21 ± 49° [I/NR]; −1 ± 6° [A/R]; −12 ± 25° [AM/NR]) for all but these three seed banks (−65, −79, and −89°).

Classification and Regression Tree Analyses

The tree model from CART analysis of actual ranges without species as an explanatory variable is shown in Figure 4. Interpretation of a tree is straightforward. The rule for each split is shown above the node, and the vertical lines leading to the child groups indicate relative importance of the split in reducing within-group variation. Each terminal group is described with the predicted spatial pattern and the seed banks assigned to that group. Misclassified seed banks have the actual pattern shown in parentheses. Structure of a tree can be interpreted by following the path down the tree to a terminal group. For example, the path to cluster I/M in Figure 4 can be interpreted as "if a seed bank had a density greater than 123 seeds m^{-2} and seeds of the weed species are small, then the I/M pattern of actual ranges of would be expected." Another interpretation, beginning with

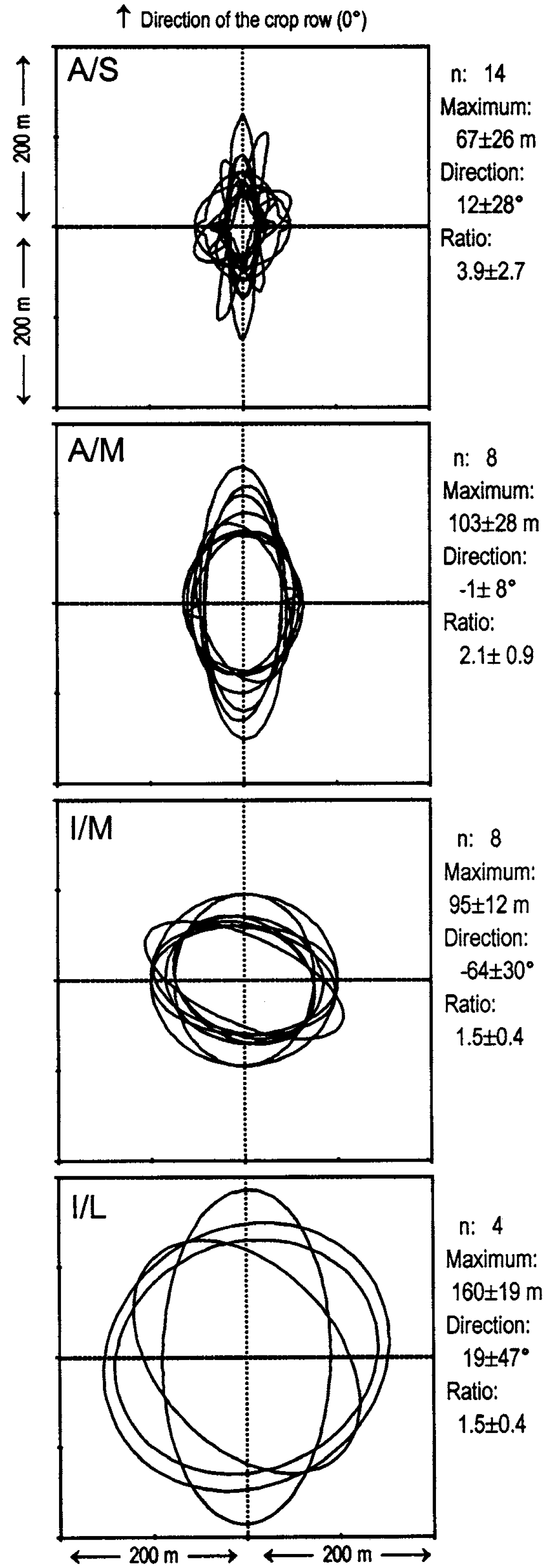

FIGURE 2. Four patterns of variation with direction of actual ranges of spatial dependence of weed seed banks in eastern Colorado. Each ellipse represents the variation in the range of spatial dependence for a seed bank of a single species, and seed banks are grouped by similar patterns of variation identified with cluster analysis. See Figure 1 for interpretation of ellipses. Data are mean value ± SD. Designations for patterns in actual ranges indicate the magnitude of the variation in the range of spatial dependence with direction (anisotropic [A] or isotropic [I]) and length of ranges (short [S], medium [M], or long [L]).

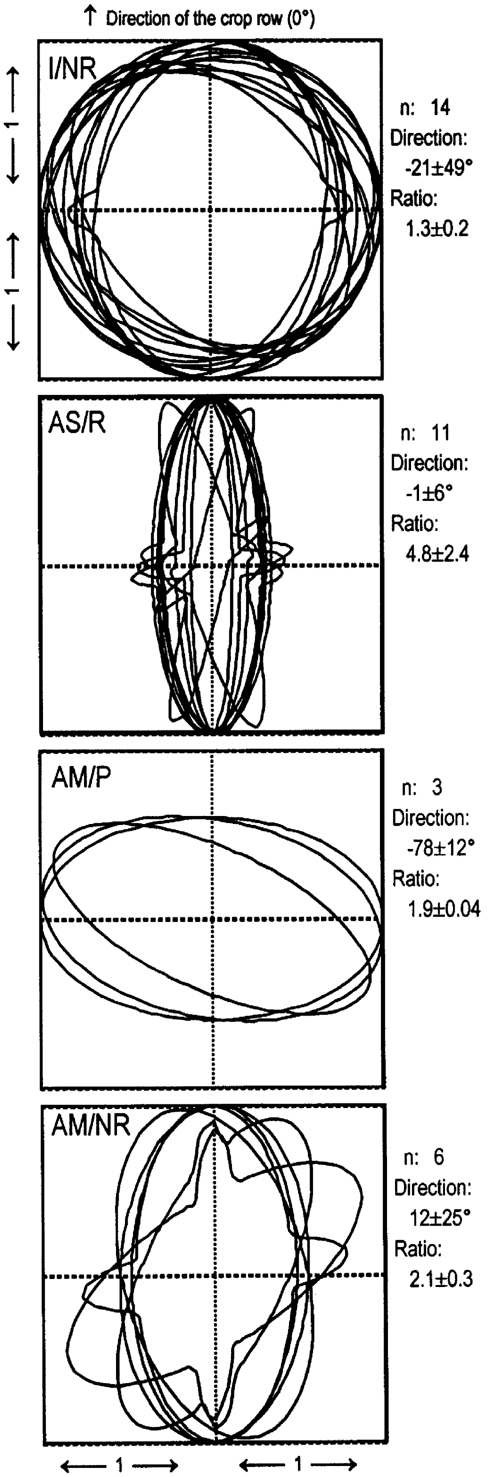

FIGURE 3. Four patterns of variation with direction of normalized ranges of spatial dependence of weed seed banks in eastern Colorado. Each ellipse represents the variation for a seed bank of a single species, and seed banks are grouped by patterns of variation identified with cluster analysis. See Figure 1 for interpretation of ellipses. Ranges were normalized by scaling to a maximum distance of one for each seed bank. Data are mean value ± SD. Designations for patterns indicate variation in the range of spatial dependence and with direction (strong anisotropy [AS], medium anisotropy [AM] or isotropic [I]), and direction of the maximum continuity (row direction [R], nearly row direction [NR] and perpendicular to row direction [P]).

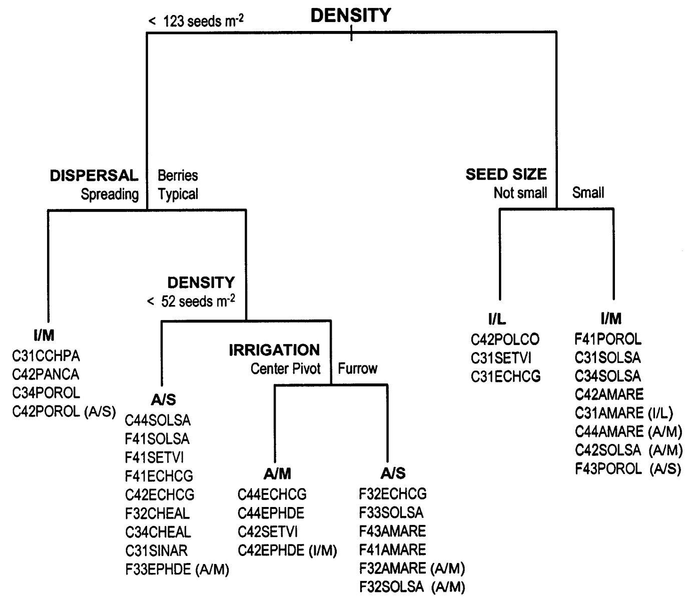

FIGURE 4. Tree model from classification and regression tree analysis to explain four patterns of variation with direction in actual ranges of spatial dependence of weed seed banks. Patterns are illustrated in Figure 2, and possible explanatory variables were demographic characteristics of species, seed bank density, and field attributes.

the terminal group, is "pattern I/M was associated with seed banks of small-seeded species that had a density greater than 123 seeds m^{-2}."

Many factors influenced the spatial distributions of the seed banks. Each model included four or five explanatory variables, and eight of the possible eleven explanatory variables appeared in at least one model (Figures 4–7). Percent organic matter, manure application, and seed production were not in any model. Misclassification rate was about 25% for all analyses, with pattern of nine seed banks misclassified in each analysis without species as an explanatory variable. With species in a model, one less seed bank was misclassified.

Consistent misclassifications in a CART analysis may indicate missing explanatory variables or, in some cases, an inadequate response variable. However, misclassification of seed banks was not consistent among the four models. Only four seed banks were misclassified in all models (C31AMARE, C42EPHDE, F32SOLSA, and F43POROL). Sixteen seed banks were misclassified in at least one model, and 18 seed banks were classified correctly in all models. The pigweed seed bank in Field C31 may have always been misclassified because manure was only applied to Field C31, and the application strongly influenced the distribution of this seed bank but not spatial pattern of the other five seed banks in the field. Maps of seed counts (not shown) support this conclusion. Misclassification was not clearly linked to any characteristic of spatial pattern (maximum range of a seed bank, ratio of anisotropy, or direction of maximum continuity) (data not shown), so incorrect grouping of seed banks (inadequate response variable) was probably not a primary cause of misclassification. Field is a reasonable candidate as an explanatory variable because this variable would capture both management practices and field attributes. However, misclassifications were nearly equally distributed among fields (data not shown).

Actual Ranges

Density was the most important factor influencing the pattern of variation in the actual ranges of spatial dependence of the seed banks. There were two splits on density in the tree model for analysis of pattern of actual ranges without species as an explanatory variable (Figure 4). Moreover, the first split (< 123 seeds m^{-2}) explained the most variability in pattern of actual ranges as indicated by the longest vertical lines leading to child nodes of any split. This split associates the most isotropic distributions and the longest ranges (I/M, I/L) with the largest seed banks. This split is consistent with the observation of Dessaint et al. (1991) that patchiness of seed banks in agricultural fields is inversely proportional to density. Density may indicate the time a seed bank has been established in a field. With time, seeds can spread farther (longer ranges) and in more directions (isotropic).

Besides density, pattern of actual ranges was explained with type of dispersal, seed size, and type of irrigation. Splits on these variables were comparable in explaining variability. The patterns associated with large seed banks (I/M, I/L) were differentiated by seed size and may reflect the influence of dispersal of seeds by wind on spatial distribution. Mortensen et al. (1993) observed that seedlings of smaller seeded species were less aggregated than species with large seeds and attributed this to greater dispersal of small seeds by wind. In this case, dispersal of small seeds by wind may explain the difference in the direction of maximum continuity between the patterns. Maximum continuity of the smaller seeded species (I/M) was primarily in the direction of prevailing wind of the area of this study compared with the more typical direction of the row for seed banks with larger seeds (I/L). Significant dispersal of seeds in the soil by wind is likely because wind erosion is a problem in sandy fields in the area of this study. Soils of the fields sampled were 37 to 88% sand.

Pattern I/M was associated with low- as well as high-density seed banks (Figure 4), but for small seed banks, this pattern was only associated with characteristics to promote dispersal over longer distances (purslane, witchgrass, and longspine sandbur). For the other small seed banks, pattern of actual ranges (A/S, A/M) was predicted with splits on density and type of irrigation. The strongest anisotropy and shortest ranges (A/S) were associated with the lowest density seed banks (< 52 seeds m^{-2}). For seed banks of intermediate density (52 to 123 seeds m^{-2}), stronger anisotropy (ratio of 3.9 ± 2.7 [A/S] compared with 2.1 ± 0.9 [A/M]) was attributed to furrow irrigation. Seeds may be moved within fields with irrigation water (Dastgheib 1989; Kelley and Bruns 1975), however, the difference in anisotropy is more likely caused by tillage that is more consistently in the direction of the crop row in furrow-irrigated fields than center pivot–irrigated fields. Tillage in center pivot–irrigated fields may spread seeds in more directions and not consistently in any direction.

One less seed bank was misclassified with species as an explanatory variable in the analysis of pattern of actual ranges, but the logic of the model was not as apparent (Figure 5). Both the number of misclassifications and the scientific soundness of the tree model should be examined because sometimes one variable may be a proxy for a different variable. The logic of the analysis of pattern of actual ranges

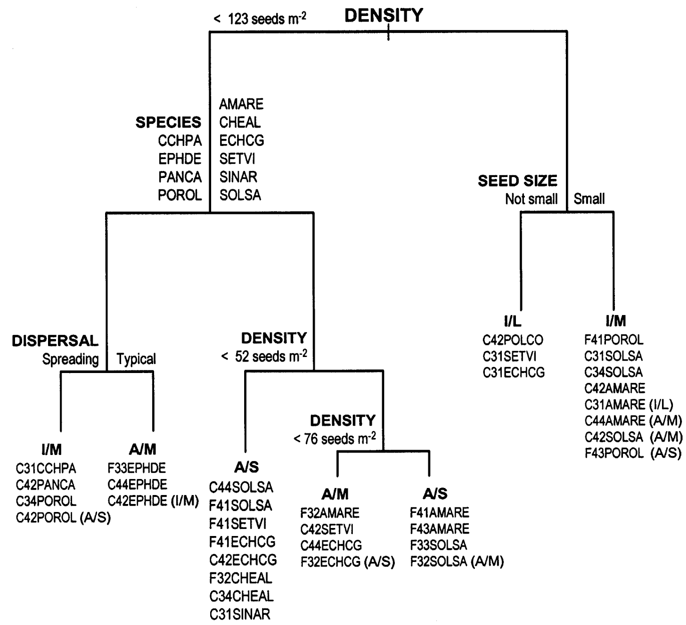

FIGURE 5. Tree model from classification and regression tree analysis to explain four patterns of variation with direction in the actual ranges of spatial dependence of weed seed banks. Patterns are illustrated in Figure 2, and possible explanatory variables were demographic characteristics of species, seed bank density, field attributes, and species.

was changed only for branches of the tree for pattern of lower density seed banks when species was added as an explanatory variable. With this change, a split on density (< 76 seeds m^{-2}) replaced a split on irrigation for lower density seed banks (< 123 seeds m^{-2}). Consequently, the tree describes the influence of seed bank density on spatial distribution at an unlikely fine scale (less than 52, 52 to 75, 76 to 122, and more than 122 seeds m^{-2}). The new split on density was probably a proxy for the effect of irrigation. With species as a explanatory variable, irrigation could not be an explanatory variable at this node with the constraint of a minimum group size of three. There were only two seed banks from center pivot fields at this node compared with four seed banks at the equivalent node in the model without species.

Normalized Ranges

Analysis of pattern of normalized ranges without species (Figure 6) only differentiated between isotropic (I/NR) and highly anisotropic distributions (AS/R). Type of irrigation explained the most variation in pattern. Patterns I/NR and AS/R appear as terminal groups for seed banks from both types of irrigation, but seed banks in furrow-irrigated fields were described as typically more anisotropic (primarily pattern AS/R) than seed banks in center pivot–irrigated fields (primarily I/NR). The indicated mechanisms leading to these patterns of normalized ranges depended on type of irrigation. In the branch for seed banks from furrow-irrigated fields, type of dispersal separated seed banks of patterns I/NR from AS/R. Isotropic distributions (I/NR) were explained by longer distance dispersal as a result of spreading growth habit (purslane) or dispersal of seeds in berries (nightshade). In center pivot–irrigated fields, longevity of seeds in the soil, type of weed, and seed bank density were

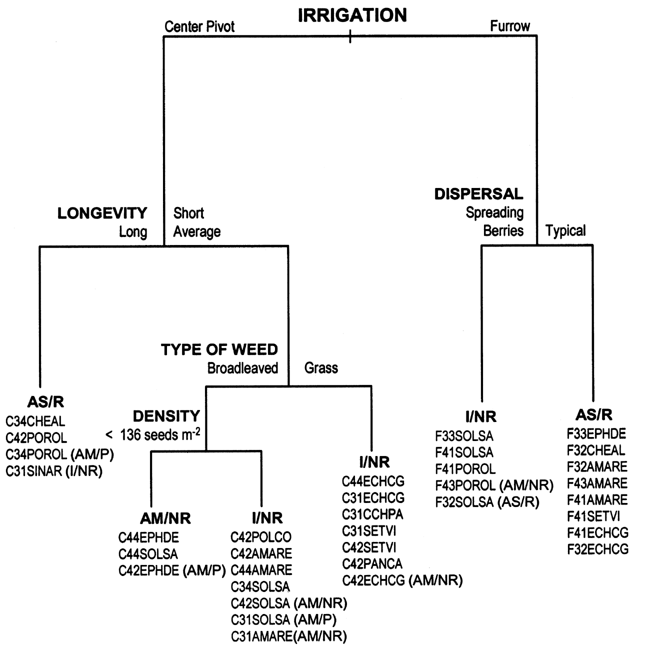

Figure 6. Tree model from classification and regression tree analysis to explain four patterns of variation in direction of normalized ranges of spatial dependence with direction of weed seed bank distributions. Patterns are illustrated in Figure 3, and possible explanatory variables were demographic characteristics of species, seed bank density, and field attributes. Ranges were normalized by scaling to a maximum distance of one for each seed bank.

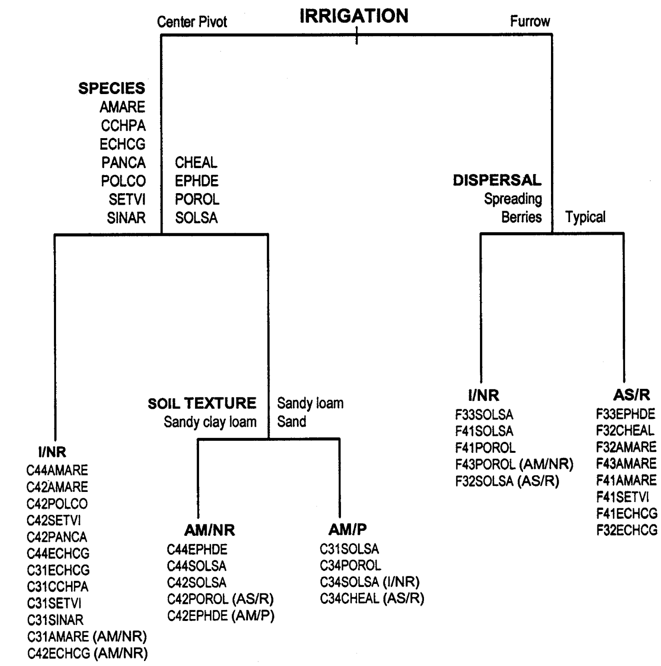

Figure 7. Tree model from classification and regression tree analysis to explain four patterns variation with direction of normalized ranges of weed seed bank distributions. Patterns are illustrated in Figure 3, and possible explanatory variables were demographic characteristics of species, seed bank density, field attributes, and species. Ranges were normalized by scaling to a maximum distance of one for each seed bank.

used to explain different patterns, but longevity was the most significant factor.

Grass seed banks were predicted to have isotropic distributions (I/NR) in center pivot–irrigated fields (Figure 6). Seed banks of broadleaved species were associated with three patterns (I/NR, AS/R, AM/NR) depending on seed longevity and seed bank density. The strongest anisotropy (AS/R) was associated with longer survival of seeds in the soil. Dessaint et al. (1991) suggest that spatial distribution of seed banks of species with short-lived seeds will be influenced primarily by natural dispersal spreading seeds in many directions because many seeds will germinate during the year after dispersal. Also, seeds with longer persistence may have a greater chance of being moved away from the original source (Rew and Cussans 1995). Pattern AS/R is consistent with spread of seeds along the crop row by field operations, whereas the isotropic pattern (I/NR) is more consistent with natural dispersal in many directions around the parent plant. Isotropic distributions were predicted for seed banks of broadleaves with short or average longevity unless the seed bank was small (AM/NR). Low density ($<$ 136 seeds m^{-2}) may indicate a recently introduced seed bank, and high density may reflect an established seed bank that has been subjected to more tillage operations. In center pivot–irrigated fields, seeds may be spread in more directions because direction of tillage may be varied, and this effect would become more apparent the longer a seed bank has been present in the field, especially with the large sampling error for seed banks. However, only half the seedbanks of pattern AM/NR were correctly classified with the split on density.

All patterns of normalized ranges were predicted, and one less seed bank was misclassified when species was included as an explanatory variable, but just as for pattern of actual ranges, the logic of the model was less credible (Figure 7). Association of pattern and factors remained the same for furrow-irrigated fields, and grasses were again predicted to have isotropic distributions in center pivot–irrigated fields. However, strong anisotropy (AS/R) was not predicted for broadleaf seed banks in center pivot–irrigated fields. The most isotropic pattern was associated with a split on species that predicted this pattern for all grasses and three of the seven broadleaf species found in center pivot fields. For the four broadleaf species with intermediate anisotropy, direction of maximum continuity was predicted with a split on soil texture. The mechanism was likely greater dispersal of seeds by wind on sandier soils. Maximum continuity was nearly in the direction of the prevailing wind of this area (AM/P; $-78 \pm 12°$) for seed banks in fields with the sandiest soils compared with the direction of the crop row (AM/NR; $-12 \pm 25°$) for the other seed banks.

Value of Comparative Studies of Seed Bank Distributions Using Classification and Regression Tree Analyses

Comparative studies can have advantages over long-term, controlled studies for investigating the spatial dynamics of seed banks. Data can be collected in 1 yr, and effects that may require several years to observe in controlled studies may be apparent. More variables can be investigated, although the values can only be partially controlled by the selection of fields. Plots do not have to be maintained, and fewer resources may be needed. Lack of appropriate methods to analyze the data has been a major obstacle to the use of comparative studies. This research indicates that CART can be valuable for identifying factors that influence the spatial distributions of seed banks and underlying mechanisms from observational data of comparative studies.

In general, explanatory variables included in the CART analyses, interactions between factors, and logic of splits were consistent with observations and inferences from previous research on the spatial dynamics of seed banks. The influence of dispersal and other processes on spatial distributions of weed seed banks depends on interacting biological and agricultural factors (Bigwood and Inouye 1988; Chauvel et al. 1989; Dessaint et al. 1991; Marshall 1989) and patchy distributions of weed populations can be created by several processes (Cousens and Woolcock 1997). In this analysis, dispersal was identified as one of the most important processes shaping spatial correlation of weed seed banks, and the type of dispersal (natural dispersal or by wind or tillage) influenced the nature of spatial correlation (Figures 4–7). The complexity of spatial dynamics of seed banks was represented by the number of explanatory variables included in a tree and multiple paths leading to terminal groups of some spatial patterns (A/S and I/M, Figures 4 and 5; I/M, Figure 4; I/NR, Figures 6 and 7; AS/R, Figure 6).

However, the CART results also suggest some new ideas about spatial dynamics of seed banks. Demographic characteristics of species and seed bank density appeared more important in shaping seed bank distributions in this research than in previous research (Bigwood and Inouye 1988; Chauvel et al. 1989; Dessaint et al. 1991). Relationships between density and spatial pattern have been observed in a few studies (Ambrosio et al. 1997; Dessaint et al. 1991, 1996) but have not been explained. The structure of the CART trees indicates that density was the most important factor in explaining pattern of actual ranges, and the influence of demographic characteristics depended on density (Figures 4 and 5). Although description of demographic characteristics in this analysis was rudimentary, each tree included at least one demographic characteristic as an explanatory variable, including the trees with species as an explanatory variable. The influence of demographic characteristics on spatial distribution of seed banks or weeds in previous studies has been inconsistent, varying from none to some influence for seed size, seed morphology, and weed phenology (Cardina et al. 1997; Cousens and Woolcock 1997). The influence of density and demographic characteristics on spatial distribution may be more predominant in this study than in previous studies because interactions were modeled. For example, the influence of density on pattern of actual ranges of seed banks in this study varied with seed size, type of dispersal, and type of irrigation (Figure 4). Also, most previous studies have been of seed banks that were recently introduced into a field and found in one small area, whereas most seed banks in this study were established as indicated by more widespread, albeit patchy distribution. Possibly, some influences may not be apparent until a seed bank is established, and conclusions about spatial dynamics of seed banks from studies of newly introduced seed banks may not apply to established seed banks.

Lack of data will be the greatest obstacle to using comparative studies and CART to understand spatial dynamics of seed banks or weed populations or other aspects of weed ecology. With a limited number of observations, combinations of values of explanatory variables may be missing, and distribution of observations between combinations may be uneven. There may be too few observations in a child or terminal group to confidently interpret a split. Important explanatory variables and interactions between variables may be missed because the number of observations of a combination may not be enough to meet the minimum size for a terminal node or other child group. Also, several splits may lead to the same child groups for a node or may lead to different child nodes but the same number of misclassified observations. Instances of these situations occurred in this study. One example is the split on density (< 76 seeds m^{-2}) for pattern of actual ranges (Figure 5) that was likely a proxy for a split on type of irrigation. In the same analysis, a split on seed production (low seed production vs. average or high seed production) would have resulted in the same terminal groups as the split on seed size. However, the split on seed production would have associated low seed production with longer distances of spatial continuity. More extensive dispersal of smaller seeds by wind seemed more reasonable as an explanation so seed size was used. Surrogate explanatory variables are not unusual with CART when the number of observations is limited. Generally, building a tree model is considered an interactive process regardless of the number of observations, and statistical packages include tools to identify and evaluate alternative tree structure (Venables and Ripley 1999). The number of observations in this study was minimal compared with the typical use of CART, nevertheless, results were reasonable. We think CART warrants further investigation as a means to more effectively use observational data to understand the complex spatial dynamics of seed banks and likely other aspects of weed ecology.

Mention of a trade name or proprietary product does not constitute a guarantee or warranty by the U.S. Department of Agriculture and does not imply approval to the exclusion of other products that may be suitable.

Sources of Materials

[1] Isaaks & Co., 1042 Wilmington Way, Redwood City, CA 94062.

Acknowledgments

This research was supported in part by USDA/CREES Grant 92-34214-7353. The authors extend appreciation to R. E. Erskine, D. H. Barlin, F. I. Zeineldin, and M. R. Collins for technical support.

Literature Cited

Ambrosio, L., J. Dorado, and J. P. Del Monte. 1997. Assessment of the sample size to estimate the weed seedbank in soil. Weed Res. 37:129–137.

Audsley, E. and S. A. Beaulah. 1996. Combining weed maps to produce a treatment for patch spraying. Asp. Appl. Biol. 46:111–117.

Benoit, D. L., N. C. Kenkel, and P. B. Cavers. 1989. Factors influencing the precision of soil seed bank estimates. Can. J. Bot. 67:2833–2840.

Bigwood, D. W. and D. W. Inouye. 1988. Spatial pattern analysis of seed banks: an improved method and optimized sampling. Ecology 69: 497–507.

Burgess, T. M., R. Webster, and A. B. McBratney. 1981. Optimal interpolation and isarithmic mapping of soil properties. IV. Sampling strategy. J. Soil Sci. 32:643–699.

Burrough, P. A. 1991. Sampling designs for quantifying map unit composition. Pages 89–125 *in* M. J. Mausbach and L. P. Wilding, eds. Spatial Variabilities of Soils and Landforms SSSA Special Publication No. 28. Madison, WI: Soil Science Society of America.

Cardina, J., G. A. Johnson, and D. H. Sparrow. 1997. The nature and consequence of weed spatial distribution. Weed Sci. 45:364–373.
Cardina, J., D. H. Sparrow, and E. McCoy. 1996. Spatial relationships between seedbank and seedling populations of common lambsquarters (*Chenopodium album*) and annual grasses. Weed Sci. 44:298–308.
Chauvel, B., J. Gasquez, and H. Darmency. 1989. Changes of weed seed bank parameters according to species, time and environment. Weed Res. 29:213–220.
Clark, L. A. and D. Pregibon. 1992. Tree-based models. Pages 377–420 *in* J. M. Chambers and T. J. Hastie, eds. Statistical Models in S. Pacific Grove, CA: Wadsworth and Brooks.
Colbach, N., F. Forcella, and G. A. Johnson. 2000. Spatial and temporal stability of weed populations over five years. Weed Sci. 48:366–377.
Cousens, R. and M. Mortimer. 1995. Dynamics of Weed Populations. Cambridge, U.K.: Cambridge University Press. Pp. 217–242.
Cousens, R. and J. L. Woolcock. 1997. Spatial dynamics of weeds: an overview. Pages 613–618 *in* Proceedings of the 1997 Brighton Crop Protection Conference—Weeds.
Dale, M.R.T. 2000. Spatial Pattern Analysis in Plant Ecology. Cambridge, U.K.: Cambridge University Press. Pp. 1–30.
Dastgheib, F. 1989. Relative importance of crop seed, manure and irrigation water as sources of weed infestations. Weed Res. 29:113–116.
De'ath, G. and K. E. Fabricus. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 8: 3178–3192.
Dessaint, F., G. Barralis, M. L. Caixinhas, J.-P. Mayor, J. Recasens, and G. Zanin. 1996. Precision of soil seed bank sampling: how many cores? Weed Res. 36:143–151.
Dessaint, F., R. Chadoeuf, and G. Barralis. 1991. Spatial pattern analysis of weed seeds in the cultivated soil seed bank. J. Appl. Ecol. 28:721–730.
Deutsch, C. V. and A. G. Journel. 1998. GSLIB Geostatistical Software Library and User's Guide. New York: Oxford University Press. Pp. 24–29.
Dieleman, J. A., D. A. Mortensen, and L. J. Young. 1999. Predicting within-field weed species occurrence based on field-site attributes. Pages 517–528 *in* J. V. Stafford, ed. Proceedings of the 2nd European Conference on Precision Agriculture; Odense, Denmark. London: SCI.
Dobbertin, M. and G. S. Biging. 1997. Using the non-parametric classifier CART to model forest tree mortality. Forest Sci. 44:507–516.
Flatman, G. T., E. J. Englund, and A. A. Yfanis. 1988. Geostatistical approaches to the design of sampling regimes. Pages 73–84 *in* L. H. Keith, ed. Principles of Environmental Sampling. Washington, DC: American Chemical Society.
Ford, E. D. and E. Renshaw. 1984. The interpretation of process from pattern using two dimensional spectral analysis: modelling single species pattern in vegetation. Vegetation 56:113–123.
Ghersa, C. M. and M. L. Roush. 1993. Searching for solutions to weed problems: do we study competition or dispersal? BioScience 43:104–109.
Gotway, C. A., R. B. Ferguson, and G. W. Hergert. 1996. The effects of mapping and scale on variable-rate fertilizer recommendations for corn. Pages 321–330 *in* P. C. Robert, R. H. Rust, and W. E. Larson, eds. Precision Agriculture. Minneapolis, MN: American Society of Agronomy.
Grundy, A. C., A. Mead, and S. Burston. 1999. Modelling the effect of cultivation on seed movement with application to the prediction of weed seedling emergence. J. Appl. Ecol. 36:663–678.
Hallahan, M. and R. Rosenthal. 2000. Interpreting and reporting results. Pages 125–131 *in* H.E.A. Tinsley and S. D. Brown, eds. Handbook of Applied Multivariate Statistics and Mathematical Modeling. San Diego, CA: Academic.
Halstead, S. J. and K. L. Gross. 1990. Geostatistical analysis of the weed seed bank. Proc. North Central Weed Sci. Soc. 45:123–124.
Hastie, T., R. Tibshirani, and J. Friedman. 2001. The Elements of Statistical Learning. New York: Springer-Verlag. Pp. 453–459.
Hausler, A. and H. Nordmeyer. 1999. Characterizing spatial and temporal dynamics of weed seedling populations. Pages 463–472 *in* Proceedings of the 2nd European Conference on Precision Agriculture; Odense, Denmark. London: SCI.
Heisel, T., S. Christensen, and A. M. Walter. 1999. Whole-field experiments with site-specific weed management. Pages 759–768 *in* J. V. Stafford, ed. Proceedings of the 2nd European Conference on Precision Agriculture; Odense, Denmark. London: SCI.
Howard, C. L., A. M. Mortimer, P. Gould, P. D. Putwain, R. Cousens, and G. W. Cussans. 1991. The dispersal of weeds: seed movement in arable agriculture. Pages 821–828 *in* 1991 Proceedings of the Brighton Crop Protection Conference—Weeds. Volume 5.
Isaaks, E. H. and R. M. Srivastava. 1989. An Introduction to Applied Geostatistics. New York: Oxford University Press. Pp. 140–182, 369–399.
Johnson, G. A., J. Cardina, and D. A. Mortensen. 1997. Site-specific weed management: current and future directions. Pages 131–147 *in* P. C. Robert, R. H. Rust, and W. E. Larson, eds. Site-specific Management for Agricultural Systems. Minneapolis, MN: American Society of Agronomy.
Jones, N. E. 1998. The number of soil cores required to accurately estimate the seed bank on arable land. Asp. Appl. Biol. 51:1–8.
Kelley, A. D. and V. F. Bruns. 1975. Dissemination of weed seeds by irrigation water. Weed Sci. 23:486–493.
Marshall, E.J.P. 1989. Distribution patterns of plants associated with arable field edges. J. Appl. Ecol. 26:247–257.
Mead, A., A. C. Grundy, and S. Burston. 1998. Predicting the movement of seeds following cultivation. Asp. Appl. Bio. 51:91–94.
Mortensen, D. A., G. A. Johnson, and L. J. Young. 1993. Weed distribution in agricultural fields. Page I/NR13–124 *in* P. C. Robert, R. H. Rust, and W. E. Larson, eds. Proceedings of Soil Specific Crop Management. Madison, WI: ASA-CSSA-SSSA.
Nordmeyer, H., A. Hausler, and P. Niemann. 1997. Patchy weed control as an approach in precision farming. Pages 307–314 *in* J. V. Stafford, ed. Precision Agriculture '97. Oxford, U.K.: Bios Scientific Publishers.
Oliver, M. A. 1999. Exploring soil spatial variation geostatistically. Pages 463–472 *in* J. V. Stafford, ed. Proceedings of the 2nd European Conference on Precision Agriculture; Odense, Denmark. London: SCI.
Oliver, M. A., Z. Frogbrook, R. Webster, and C. J. Dawson. 1997. A rational strategy for determining the number of cores for bulked soil samples. Pages 155–162 *in* J. V. Stafford, ed. Precision Agriculture '97. Oxford, U.K.: Bios Scientific.
Paice, M.E.R., W. Day, L. J. Rew, and A. Howard. 1998. A stochastic simulation model for evaluating the concept of patch spraying. Weed Res. 38:373–388.
Rew, L. J. and R. D. Cousens. 2001. Spatial distribution of weeds in arable crops: are current sampling and analytical methods appropriate? Weed Res. 41:1–18.
Rew, L. J. and G. W. Cussans. 1995. Patch ecology and dynamics: how much do we know? Pages 1059–1068 *in* Proceedings of the 1995 Brighton Crop Protection Conference—Weeds.
Rew, L. J. and G. W. Cussans. 1997. Horizontal movement of seeds following tine and plough cultivation: implications for spatial dynamics of weed infestations. Weed Res. 37:247–256.
Rossi, R. R., D. J. Mulla, A. G. Journel, and E. H. Franz. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecol. Monogr. 62:277–314.
[SAGE2001] Spatial and Geostatistical Environment for Variography. 1999. User's Guide. San Mateo, CA: Isaaks. Pp. 27–40.
[SAS] Statistical Analysis Systems. 1988. SAS/STAT User's Guide. Release 6, 3rd ed. Cary, NC: Statistical Analysis Systems Institute. Pp. 283–357.
Venables, W. N. and B. D. Ripley. 1999. Modern Applied Statistics with S-Plus. New York: Springer-Verlag. Pp. 303–319.
Walter, A. M., T. Heisel, and S. Christensen. 1997. Shotcuts in weed mapping. Pages 777–784 *in* J. V. Stafford, ed. Precision Agriculture '97. Oxford, U.K.: Bios Scientific Publishers.
Weisz, R., S. Fleischer, and Z. Smilowitz. 1995. Site-specific integrated pest management for high value crops: sample units for map generation using the Colorado potato beetle (Coleoptera: Chrysomelidae) as a model system. J. Econ. Entom. 88:1069–1080.
Wiles, L. J., D. H. Barlin, E. E. Schweizer, H. R. Duke, and D. E. Whitt. 1996. A new soil sampler and elutriator for collecting and extracting weed seeds from soil. Weed Technol. 10:35–41.
Wiles, L. J. and E. E. Schweizer. 2002. Spatial dependence of weed seed banks and strategies for sampling. Weed Sci. 50:595–606.
Wilson, B. J. and P. Brain. 1991. Long-term stability of distribution of *Alopecurus myosuroides* Huds. within cereal fields. Weed Res. 31:367–373.

Received May 8, 2003, and approved July 3, 2004.